Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Multivariate Analysis 97 (2006) 720–732
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Consistency of the generalized MLE of a joint distribution function with multivariate interval-censored data

Shaohua Yu[a, b], Qiqing Yu[a, b, *], George Y.C. Wong[a, b]

[a] *Novartis Pharmaceuticals Corporation, New Jersey, Binghamton University, New York, USA*
[b] *Strang Cancer Prevention Center, New York, USA*

Received 22 September 2003
Available online 1 September 2005

Abstract

Wong and Yu [Generalized MLE of a joint distribution function with multivariate interval-censored data, J. Multivariate Anal. 69 (1999) 155–166] discussed generalized maximum likelihood estimation of the joint distribution function of a multivariate random vector whose coordinates are subject to interval censoring. They established uniform consistency of the generalized MLE (GMLE) of the distribution function under the assumption that the random vector is independent of the censoring vector and that both of the vector distributions are discrete. We relax these assumptions and establish consistency results of the GMLE under a multivariate mixed case interval censorship model. van der Vaart and Wellner [Preservation theorems for Glivenko–Cantelli and uniform Glivenko–Cantelli class, in: E. Gine, D.M. Mason, J.A. Wellner (Eds.), High Dimensional Probability, vol. II, Birkhäuser, Boston, 2000, pp. 115–133] and Yu [Consistency of the generalized MLE with multivariate mixed case interval-censored data, Ph.D Dissertation, Binghamton University, 2000] independently proved strong consistency of the GMLE in the $L_1(\mu)$-topology, where μ is a measure derived from the joint distribution of the censoring variables. We establish strong consistency of the GMLE in the topologies of weak convergence and pointwise convergence,

* Corresponding author. Department of Mathematical Sciences, Binghamton University, Binghamton, NY 13902, USA. Fax: +607 777 2450.

E-mail address: qyu@math.binghamton.edu (Qiqing Yu).

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2005.07.006

and eventually uniform convergence under appropriate distributional assumptions and regularity conditions.

AMS 1991 subject classification: primary 62G20; secondary 62H12

Keywords: General maximum likelihood estimation; Consistency; Multivariate interval censorship model; Mixed case model; Case k model

1. Introduction

Interval-censored (IC) data arise naturally in longitudinal follow-up studies in which the exact value of a time-to-event variable X cannot be directly observed and is instead known to lie in an observable interval defined by two successive follow-up or inspection time points. Examples of such an X variable include time to relapse of a disease such as cancer, and time for the value of a surrogate endpoint biomarker to reach a target value in a chemoprevention intervention trial. Denote the observable interval by $I = (L, R]$, where L and R are extended random variables such that $-\infty \leqslant L < X \leqslant R \leqslant \infty$. The statistical question of interest is generalized maximum likelihood (GML) estimation of the distribution function $F(x) = Pr(X \leqslant x)$, or equivalently, the survival function $S(x) = 1 - F(x)$, under a specified model for the IC data.

The simplest model for IC data is the case 1 model (data from the model are also called current status data, see [1]) in which there is only one inspection time Y, which is independent of X. If the event has taken place by the inspection time so that $X \leqslant Y$, then $(L, R) = (-\infty, Y)$; otherwise, $(L, R) = (Y, \infty)$.

The case 2 model (see [4]) is another model for IC data in which there are two censoring random variables $U < V$. The event can take place before U, between U and V, or has not occurred by the time V. The corresponding $(L, R]$ intervals are $(-\infty, U]$, $(U, V]$ or $(V, \infty]$, respectively. In a longitudinal follow-up study involving IC data, a subject is evaluated at random successive inspection times $Y_1, \ldots, Y_K$, where $K \geqslant 2$ is the number of inspection times. The relation between the random vector (U, V) and the inspection times $Y_1, \ldots, Y_K$, are as follows. If X is neither right censored nor left censored, take (U, V) to be the two successive inspection time points (Y_i, Y_{i+1}) such that $Y_i < X \leqslant Y_{i+1}$. If X is left censored, take $U = Y_1$ and $V = Y_2$. If X is right censored, take $V = Y_K$ and $U = Y_{K-1}$. The random variables U and V thus defined are not independent of X, even when $Y_1, \ldots, Y_K$ are independent of X, unless $K = 2$ with probability one. However, for the convenience of the proof of asymptotic properties of the GMLE of F_0, it is often assumed that (U, V) and X are independent (see [4,19]). In doing so, it is implicitly assumed that $(U, V, K) = (Y_1, Y_2, 2)$ with probability one.

Wellner [14] generalized the case 1, 2 models to a case k model in which there are $K(\geqslant 1)$ inspection times $Y_1 < \cdots < Y_K$, where K is fixed for each X in a random sample. The observable intervals $(L, R]$ consist of $(-\infty, Y_1]$, $(Y_1, Y_2]$, $\ldots$, $(Y_{k-1}, Y_K]$,$(Y_K, \infty]$ that contain X. As in case 2 model, it is often assumed that $Y_1, \ldots, Y_K$ are independent of X.

In a typical longitudinal follow-up study involving IC data, the number of inspection times K is random. If one assumes that K is fixed and $Y_1, \ldots, Y_K$ are independent of X, then none of case K models, $K \geqslant 2$, is appropriate for such data.

Schick and Yu [8] proposed a model for such IC data which allows the number of inspection times to be random and assume that inspection times are independent of X. They called their model the mixed case model because it can be viewed as a mixture of various case k models. They established strong consistency of the GMLE under the model. Their result implies that the independent assumption between (U, V) and X in the case 2 model can be replaced by a weaker assumption in the proof of asymptotic properties of the GMLE of F_0. The mixed case model has been adopted by Wellner and Zhang [15], van der Vaart and Wellner [13], Zhang et al. [21], Sun and Fand [11], Ren [6] and Song [10].

Multivariate interval censoring involves $d \geqslant 2$ correlated variables $X_1, \ldots, X_d$, each of which is subject to interval censoring. A multivariate IC observation consists of d pairs of observations (L_δ, R_δ), where $0 \leqslant L_\delta < X_\delta \leqslant R_\delta \leqslant \infty$, $\delta = 1, \ldots, d$. GML estimation of the joint distribution function $F_0(\mathbf{x}) = Pr(X_1 \leqslant x_1, \ldots, X_d \leqslant x_d)$ of $\mathbf{X}$ is of interest, where $\mathbf{x} = (x_1, \ldots, x_d)'$ and $\mathbf{X} = (X_1, \ldots, X_d)'$. An illustrative example of bivariate interval censoring is provided by a case–control cataract study reported by the Italian–American Cataract Study Group [12]. In the study, X_1 and X_2 refer to time to progression of cortical, nuclear, or posterior subcapsular cataracts of the left and the right eyes of a patient, respectively.

Wong and Yu [16] discussed GML estimation of F_0 under a multivariate case 2 IC model. They established uniform consistency and derived asymptotic properties of the GMLE under discrete distributional assumptions. A multivariate mixed case IC model was considered in Example 1 of van der Vaart and Wellner [13] and investigated in a Ph.D dissertation by Yu [17]. They independently proved strong consistency of the GMLE of F_0 in the $L_1(\mu)$-topology, where μ is a measure derived from the joint distribution function of the inspection times. Song [9] established several consistency results in other topologies under the case 1 interval censorship model. However, strong consistency of the GMLE in other topologies under the mixed case model has not been reported in the literature.

In Section 2, we formulate the multivariate mixed case IC model and present the consistency result in the $L_1(\mu)$-topology. We present strong consistency results in other topologies in Section 3. Details of some proofs are relegated to Section 4.

This paper represents the multivariate extension of Schick and Yu [8]. As expected, the generalization from the univariate case to the multivariate case is not straightforward. For instance, while the GMLE-induced measure of each maximum intersection of the observed intervals is unique in the univariate interval censoring, it is no longer so in the multivariate case [20]. A key in the consistency proof in the univariate mixed case model is Helly's Selection Theorem (see [7]), which guarantees the pointwise convergence of a subsequence of distribution functions on $\mathbb{R}$. However, in higher dimensions $\mathbb{R}^d$ ($d > 1$), Helly's Selection Theorem [2] only gives pointwise convergence on continuity points of the limiting function. Thus, topology of pointwise convergence on $\mathbb{R}^d$ is not valid. We find an approach to bypass this difficulty.

For univariate right-censored data, uniform convergence of the GMLE of a distribution function F_0 is a direct consequence of continuity of F_0 (see [18]). Influenced by such a result, Gentleman and Geyer [3] and Huang [5] claimed uniform consistency results for the GMLE with univariate IC data. As pointed out by Schick and Yu [8], both of their theorems are false and the GMLE cannot be uniformly consistent under the conditions stated in their theorems. Thus our proofs of uniform strong consistency results are not trivial.

2. Notations and preliminary results

Let $\mathbf{K} = (K_1, \ldots, K_d)'$ be a vector of positive random integers, where K_i stands for the total number of inspection times corresponding to X_i, $i = 1, \ldots, d$. Throughout the paper, we assume that $E(\prod_{i=1}^d K_i) < \infty$. This is a mild assumption and is generally satisfied in practice.

The *multivariate mixed case* model is formulated as follows. Conditional on $\mathbf{K} = (k_1, \ldots, k_d)'$, let the random vector $\mathbf{Y} = \{Y_{\delta,k_\delta,j} : \delta = 1, \ldots, d \text{ and } j = 1, \ldots, k_\delta\}$, where $k_\delta \in \mathbb{Z}^+$ (the set of all positive integers) and $Y_{\delta,k_\delta,1} < \cdots < Y_{\delta,k_\delta,k_\delta}$ are random inspection times for the δth coordinate. Assume that $(\mathbf{K}, \mathbf{Y})$ and $\mathbf{X}$ are independent. On the event $\{\mathbf{K} = (k_1, \ldots, k_d)'\}$, let $(\mathbf{L}, \mathbf{R}) = (L_1, R_1, \ldots, L_d, R_d)$ such that each pair (L_δ, R_δ) is from a univariate mixed case model, i.e., (L_δ, R_δ) denotes the endpoints of the random interval among

$$(-\infty, Y_{\delta,k_\delta,1}], \quad (Y_{\delta,k_\delta,1}, Y_{\delta,k_\delta,2}], \ldots, (Y_{\delta,k_\delta,k_\delta-1}, Y_{\delta,k_\delta,k_\delta}], \quad (Y_{\delta,k_\delta,k_\delta}, \infty)$$

that contains X_δ, where $Y_{\delta,k_\delta,0} = -\infty$ and $Y_{\delta,k_\delta,k_\delta+1} = \infty$, $k_\delta \in \mathbb{Z}^+$. Let $(\mathbf{L}_1, \mathbf{R}_1), \ldots, (\mathbf{L}_n, \mathbf{R}_n)$ be independent copies of the pair of $(\mathbf{L}, \mathbf{R})$. Define the generalized likelihood function as follows:

$$\Lambda_n(F) = \prod_{\eta=1}^{n} \mu_F((L_{\eta,1}, R_{\eta,1}] \times \cdots \times (L_{\eta,d}, R_{\eta,d}]),$$

where F is a distribution function and μ_F is the measure on $\mathbb{R}^d$ induced by F. We call a maximizer $\hat{F}_n$ of Λ_n a GMLE of F_0. The GMLE of F_0 can be obtained by the same algorithm for obtaining the GMLE under the multivariate case 2 IC model (see [16]), because the likelihood function $\Lambda(F)$ only depends on the sufficient statistics $(\mathbf{L_1}, \mathbf{R_1}), \ldots, (\mathbf{L_n}, \mathbf{R_n})$, which are of the same form under both models. Yu [20] discussed how to address the issue that the GMLE is not unique. We refer the technical details to their papers. Moreover, under the assumptions in this paper, the non-uniqueness of the GMLE would not affect the consistency results.

Let $\mathcal{M}$ be the collection of all intervals in $\mathbb{R}$. Let $\mathcal{W}$ be the collection of all finite unions of rectangles, $A_1 \times \cdots \times A_d$, where $A_1, \ldots, A_d \in \mathcal{M}$. Obviously $\mathcal{W}$ is an algebra. Since we consider the limit of the GMLE, and the limit of a sequence of distribution functions may not be a proper distribution function, we shall define a collection $\mathcal{F}$ so that each limit of a sequence of distribution functions belongs to $\mathcal{F}$. For this purpose, we first define $F(\mathbf{x}) = 0$ if one of the coordinates of $\mathbf{x}$ is $-\infty$ and $F(\infty, \ldots, \infty) = 1$ for each distribution function F. It can be shown that an $F \in \mathcal{F}$ is a function from $\overline{\mathbb{R}^d}$ into [0, 1] such that

1. F is nondecreasing in each variable;

2. $\mu_F(W) \geqslant 0$ for each $W \in \mathcal{W}$;
3. $F(\infty, \ldots, \infty) = 1$ and $F(\mathbf{x}) = 0$ if one of the coordinates of $\mathbf{x}$ is $-\infty$.

In this paragraph, we illustrate the above notation in the bivariate case ($d = 2$). Then, $\mathbf{K} = \binom{K_1}{K_2}$, $\mathbf{X} = \binom{X_1}{X_2}$, $\mathbf{Y} = \binom{\mathbf{Y}_1}{\mathbf{Y}_2}$, where

$$\mathbf{Y}_\delta = \begin{pmatrix} Y_{\delta,1,1} & & \\ Y_{\delta,2,1} & Y_{\delta,2,2} & \\ Y_{\delta,3,1} & Y_{\delta,3,2} & Y_{\delta,3,3} \\ \ldots & & \end{pmatrix}$$

for each $\delta = 1, 2$. $(\mathbf{L}_\eta, \mathbf{R}_\eta) = (L_{\eta,1}, R_{\eta,1}, L_{\eta,2}, R_{\eta,2}), \eta = 1, \ldots, n$. A set function induced by some function $F \in \mathcal{F}$, say μ_F, restricted on $\mathcal{W}$ is

$$\mu_F(W) = \begin{cases} F(b,d) + F(a,c) - F(a,d) - F(b,c) & \text{if } W = (a,b] \times (c,d], \\ F(a,d) + F(a-,c) - F(a-,d) - F(a,c) & \text{if } W = [a,a] \times (c,d], \\ F(b,c) + F(a,c-) - F(b,c-) - F(a,c) & \text{if } W = (a,b] \times [c,c], \\ F(a,c) + F(a-,c-) - F(a,c-) - F(a-,c) & \text{if } W = [a,a] \times [c,c], \end{cases}$$

where $F(\mathbf{x}-) = \sup\{F(\mathbf{t}) : \mathbf{t} < \mathbf{x}\}$, $F(a, c-) = \sup\{F(a,t) : t < c\}$ and $F(a-, c) = \sup\{F(t,c) : t < a\}$. Also, the notion $\mathbf{x} \leqslant \mathbf{y}$ [$\mathbf{x} < \mathbf{y}$] means $x_i \leqslant y_i$ [$x_i < y_i$], for all $i = 1, 2$. The measure induced by a distribution function F is such a set function.

Define a measure μ on the Borel σ-field $\mathcal{B}(\mathbb{R}^d)$ such that for each $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\begin{aligned} \mu(B) = & \sum_{k_1=1}^{\infty} \cdots \sum_{k_d=1}^{\infty} P\Big\{\mathbf{K} = \mathbf{k}\Big\} \cdot \sum_{i_1=1}^{k_1} \cdots \\ & \times \sum_{i_d=1}^{k_d} P\Big\{(Y_{1,k_1,i_1}, \ldots, Y_{d,k_d,i_d})' \in B \,\Big|\, \mathbf{K} = \mathbf{k}\Big\}, \end{aligned}$$

where $\mathbf{k} = (k_1, \ldots, k_d)'$.

Strong consistency in $L_1(\mu)$-topology is established in the theorem below.

Theorem 2.1. $\int |\hat{F}_n - F_0| d\mu \to 0$ *a.s.*

Yu [17] gave a proof of Theorem 2.1. The proof depends on the key observation that μ is a finite measure. This is because for each $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\mu(B) \leqslant \sum_{k_1=1}^{\infty} \cdots \sum_{k_d=1}^{\infty} k_1 \times \cdots \times k_d \cdot P\Big\{\mathbf{K} = \mathbf{k}\Big\} = E(K_1 \times \cdots \times K_d) < \infty$$

by our assumption. A proof of Theorem 2.1 was also given independently by van der Vaart and Wellner [13] as a corollary of their Theorems 9 and 10. Their proof is quite different from the direct proof given in [17].

Pointwise convergence for each μ-positive inspection time can be obtained as a consequence of Theorem 2.1 since $\mu(\{\mathbf{a}\})|\hat{F}_n(\mathbf{a}) - F_0(\mathbf{a})| \leqslant \int |\hat{F}_n - F_0|\, d\mu$ for each $\mathbf{a} \in \mathbb{R}^d$. We state this result in Corollary 2.1.

Corollary 2.1. $\hat{F}_n(\mathbf{a}) \to F_0(\mathbf{a})$ *a.s. for each* $\mathbf{a}$ *that satisfies* $\mu(\{\mathbf{a}\}) > 0$.

3. Propositions

Strong consistency in the topologies of weak convergence, pointwise convergence and uniform convergence are established in this section as a consequence of Theorem 2.1 with additional assumptions.

For simplicity and without loss of generality, in the proof of this section, we restrict our attention to the bivariate case ($d = 2$). However, whenever there is no confusion arisen, we use d instead of $d = 2$ so that the notation is applicable for $d \geqslant 2$.

Let $\mathbf{a}, \mathbf{b}, \mathbf{x}$ be members of $\mathbb{R}^d$. For convenience, we adopt the following notations:

$$\begin{aligned}
&\big[\mathbf{a}, \mathbf{b}\big] = [a_1, b_1] \times \cdots \times [a_d, b_d], \quad \mathbf{a} \leqslant \mathbf{b},\\
&\big[\mathbf{a}, \mathbf{b}\big) = [a_1, b_1) \times \cdots \times [a_d, b_d) \quad \text{and} \quad \big(\mathbf{a}, \mathbf{b}\big] = (a_1, b_1] \times \cdots \times (a_d, b_d], \ \mathbf{a} < \mathbf{b},
\end{aligned}$$

$(\mathbf{a}, \mathbf{b})$ is the interior set of $[a_1, b_1] \times \cdots \times [a_d, b_d]$, for instance, for $d = 2$,

$$\big(\mathbf{a}, \mathbf{b}\big) = \begin{cases} (a_1, b_1) \times (a_2, b_2) & \text{if } \mathbf{a} < \mathbf{b}, \\ [a_1, a_1] \times (a_2, b_2) & \text{if } a_1 = b_1 \quad \text{and} \quad a_2 < b_2, \\ (a_1, b_1) \times [a_2, a_2] & \text{if } a_2 = b_2 \quad \text{and} \quad a_1 < b_1. \end{cases}$$

We say that F is *continuous from above* at $\mathbf{x}$, if for each $\varepsilon > 0$, there exists a $\delta > 0$ such that $\mathbf{x} \leqslant \mathbf{y} < \mathbf{x} + \delta\mathbf{1}$ ($\mathbf{1}$ is the unit vector) implies that $|F(\mathbf{y}) - F(\mathbf{x})| < \varepsilon$. We define $\mathbf{x}$ to be *a support point* of μ, if $\mu((\mathbf{x} - \delta\mathbf{1}, \mathbf{x}] \cup [\mathbf{x}, \mathbf{x} + \delta\mathbf{1})) > 0$ for all $\delta > 0$. Let $\mathcal{S}_\mu$ denote the set of all support points of μ. Define $\mathbf{x}$ to be a *regular point* of μ if $\mu((\mathbf{x} - \delta\mathbf{1}, \mathbf{x}]) > 0$ and $\mu([\mathbf{x}, \mathbf{x} + \delta\mathbf{1})) > 0$ for all $\delta > 0$. We say $\mathbf{x}$ is *strongly regular* with respect to μ if $\mathbf{x}$ is a regular point of μ and $\mu((\mathbf{x} - \delta\mathbf{1}, \mathbf{x})) > 0$ for all $\delta > 0$. We say that *F is continuous on a set E* if $\forall\, \mathbf{x} \in E$ and $\forall\, \varepsilon > 0$, there exits a $\delta > 0$ such that $|F(\mathbf{y}) - F(\mathbf{x})| < \varepsilon$ for all $\mathbf{y} \in E$ with $\rho(\mathbf{x}, \mathbf{y}) < \delta$. Here $\rho(\mathbf{x}, \mathbf{y}) = (\sum_{i=1}^d (x_i - y_i)^2)^{\frac{1}{2}}$. Let $\mathcal{C}_{F_0}$ denote the set of all continuity points of F_0. For convenience, we say F is *monotone* if a bounded function F is nondecreasing in each coordinate. Finally, we let $\mathcal{I}_{F_0}$ denote the set of points on which F_0 is strictly increasing, i.e., for each $\mathbf{x} \in \mathcal{I}_{F_0}$ and for all $\delta > 0$, $F_0(\mathbf{x} + \delta\mathbf{1}) > F_0(\mathbf{x} - \delta\mathbf{1})$. Now, consider $\Omega_\mu = \big\{\omega : \int_{\mathbb{R}^d} |\hat{F}_n(\mathbf{x}; \omega) - F_0(\mathbf{x})| d\mu(\mathbf{x}) \to 0 \text{ as } n \to \infty\big\}$. By Theorem 2.1, $P\{\Omega_\mu\} = 1$.

The following notations and definitions are needed in the regularity conditions on the boundary sets. For simplicity, we only make use of the notations in the bivariate case.

$\partial_l\big[\mathbf{a}, \mathbf{b}\big]$ is the left vertical boundary $[a_1, a_1] \times [a_2, b_2]$,
$\partial_r\big[\mathbf{a}, \mathbf{b}\big]$ is the right vertical boundary $[b_1, b_1] \times [a_2, b_2]$,

$\partial_u\big[\mathbf{a},\mathbf{b}\big]$ is the upper horizontal boundary $[a_1, b_1] \times [b_2, b_2]$,
$\partial_b\big[\mathbf{a},\mathbf{b}\big]$ is the bottom horizontal boundary $[a_1, b_1] \times [a_2, a_2]$,
$\partial\big[\mathbf{a},\mathbf{b}\big] = \partial_l\big[\mathbf{a},\mathbf{b}\big] \cup \partial_r\big[\mathbf{a},\mathbf{b}\big] \cup \partial_u\big[\mathbf{a},\mathbf{b}\big] \cup \partial_b\big[\mathbf{a},\mathbf{b}\big]$;
$\mathcal{Q}$ is a square, $\rule[0.1ex]{1em}{0.15ex}$ is a horizontal line segment and $\rule{0.15ex}{1.2em}$ is a vertical line segment.

Let Ψ denote $\mathcal{Q}$, $\rule[0.1ex]{1em}{0.15ex}$ or $\rule{0.15ex}{1.2em}$. Define $\mathbf{1}_\Psi = \begin{cases} (1,1)' & \text{if } \Psi = \mathcal{Q}, \\ (1,0)' & \text{if } \Psi = \rule[0.1ex]{1em}{0.15ex}, \\ (0,1)' & \text{if } \Psi = \rule{0.15ex}{1.2em}\,. \end{cases}$

For each $\delta > 0$, define $\Psi_\delta(\mathbf{x}) = \big(\mathbf{x}, \mathbf{x} + \delta\mathbf{1}_\Psi\big)$, $\Psi_{-\delta}(\mathbf{x}) = \big(\mathbf{x} - \delta\mathbf{1}_\Psi, \mathbf{x}\big)$,
$\Psi_\delta[\mathbf{x}) = \big[\mathbf{x}, \mathbf{x} + \delta\mathbf{1}_\Psi\big)$, $\Psi_{-\delta}(\mathbf{x}] = \big(\mathbf{x} - \delta\mathbf{1}_\Psi, \mathbf{x}\big]$.

Finally, let $\mathcal{G}_\delta(\mathbf{x})$, $-\!\!+\!\!-_\delta(\mathbf{x})$ and $\dagger_\delta(\mathbf{x})$denote the unions $\Psi_{-\delta}(\mathbf{x}] \cup \Psi_\delta[\mathbf{x})$, for $\Psi = \mathcal{Q}$, $\rule[0.1ex]{1em}{0.15ex}$ and$\rule{0.15ex}{1.2em}$, respectively.

We call $\mathbf{x}$ *a horizontal support point of* μ if $\mu(-\!\!+\!\!-_\delta(\mathbf{x})) > 0$ for all $\delta > 0$. Let $\mathcal{S}1_\mu$ denote the set of all horizontal support points of μ. Similarly, we define $\mathbf{x}$ to be *a vertical support point of* μ, if $\mu(\dagger_\delta(\mathbf{x})) > 0$ for all $\delta > 0$, and let $\mathcal{S}2_\mu$ be the set of all vertical support points of μ. We say $\mathbf{x}$ *a horizontal* [*vertical*] *continuity point of* F if for each $\varepsilon > 0$ there is a $\delta > 0$ such that $|F(\mathbf{y}) - F(\mathbf{x})| < \varepsilon$ for all $\mathbf{y} \in -\!\!+\!\!-_\delta(\mathbf{x})$ [$\dagger_\delta(\mathbf{x})$]. Let $\mathcal{C}1_{F_0}$ [$\mathcal{C}2_{F_0}$] denote the set of all horizontal [vertical] continuity points of F_0.

Strong consistency of $\hat{F}_n$ on the set of all regular continuity points is given by the first proposition.

Proposition 3.1. *Suppose* $\mathbf{x} \in \mathcal{C}_{F_0}$ *is a regular point of* μ, *then* $\hat{F}_n(\mathbf{x};\omega) \to F_0(\mathbf{x})$, *for each* $\omega \in \Omega_\mu$.

The next proposition gives weak convergence of $\hat{F}_n$ on the set of continuity points of F_0 on an open rectangle or an open line segment.

Proposition 3.2. *Let* $\mathbf{a} \leqslant \mathbf{b}$ *and* $\mathbf{a} \neq \mathbf{b}$, $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$. *Then the following assertions hold*:

(i) $a_1 = b_1$ *and* $\big(\mathbf{a},\mathbf{b}\big) \subset \mathcal{S}2_\mu$ *imply that for each* $\omega \in \Omega_\mu$,
$\hat{F}_n(\mathbf{x};\omega) \to F_0(\mathbf{x})$ *for all* $\mathbf{x} \in \big(\mathbf{a},\mathbf{b}\big) \cap \mathcal{C}2_{F_0}$;
(ii) $a_2 = b_2$ *and* $\big(\mathbf{a},\mathbf{b}\big) \subset \mathcal{S}1_\mu$ *imply that for each* $\omega \in \Omega_\mu$,
$\hat{F}_n(\mathbf{x};\omega) \to F_0(\mathbf{x})$ *for all* $\mathbf{x} \in \big(\mathbf{a},\mathbf{b}\big) \cap \mathcal{C}1_{F_0}$;
(iii) $\mathbf{a} < \mathbf{b}$ *and* $\big(\mathbf{a},\mathbf{b}\big) \subset \mathcal{S}_\mu$ *imply that for each* $\omega \in \Omega_\mu$,
$\hat{F}_n(\mathbf{x};\omega) \to F_0(\mathbf{x})$ *for all* $\mathbf{x} \in \big(\mathbf{a},\mathbf{b}\big) \cap \mathcal{C}_{F_0}$.

Weak convergence properties of $\hat{F}_n$ follow directly from Proposition 3.2.

Proposition 3.3. *Suppose* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ *satisfy that* $F_0(\mathbf{a}) = 0$, $F_0(\mathbf{b}-) = 1$ *and* $\mu_{F_0}(\big[\mathbf{a},\mathbf{b}\big]) = 1$. *Then the following assertions hold*:

(i) $a_1 = b_1$ *and* $\big(\mathbf{a},\mathbf{b}\big) \subset \mathcal{S}2_\mu$ *imply that for each* $\omega \in \Omega_\mu$,
$\hat{F}_n(\mathbf{x};\omega) \to F_0(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{C}2_{F_0}$;

(ii) $a_2 = b_2$ *and* $\big(\mathbf{a}, \mathbf{b}\big) \subset \mathcal{S}1_\mu$ *imply that for each* $\omega \in \Omega_\mu$, $\hat{F}_n(\mathbf{x}; \omega) \to F_0(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{C}1_{F_0}$;
(iii) $\mathbf{a} < \mathbf{b}$, $\partial_b\big[\mathbf{a}, \mathbf{b}\big] \cup \partial_u\big[\mathbf{a}, \mathbf{b}\big] \subset \mathcal{S}1_\mu$, $\partial_l\big[\mathbf{a}, \mathbf{b}\big] \cup \partial_r\big[\mathbf{a}, \mathbf{b}\big] \subset \mathcal{S}2_\mu$ *and* $\big(\mathbf{a}, \mathbf{b}\big) \subset \mathcal{S}_\mu$ *imply that for each* $\omega \in \Omega_\mu$, $\hat{F}_n(\mathbf{x}; \omega) \to F_0(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{C}_{F_0}$.

Proposition 3.4. *If every* $\mathbf{y} \in \mathcal{I}_{F_0}$ *is strongly regular with respect to* μ, *then for each* $\omega \in \Omega_\mu$, $\hat{F}_n(\mathbf{x}; \omega) \to F_0(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{C}_{F_0}$.

Pointwise convergence on an open rectangle or an open line segment follows from Theorem 2.1 and Proposition 3.2. Similarly, pointwise convergence on the entire $\mathbb{R}^2$ plane follows from Theorem 2.1 and Propositions 3.3 and 3.4.

Corollary 3.1. *Let* $\mathbf{a} \leqslant \mathbf{b}$ *and* $\mathbf{a} \neq \mathbf{b}$. *Suppose one of the assumptions listed in Proposition* 3.2 *is satisfied and* $\mu(\{\mathbf{y}\}) > 0$ *for each* $\mathbf{y} \in \big(\mathbf{a}, \mathbf{b}\big) \backslash \mathcal{C}_{F_0}$. *Then for each* $\omega \in \Omega_\mu$, $\hat{F}_n(\mathbf{x}; \omega) \to F_0(\mathbf{x})$ *for all* $\mathbf{x} \in \big(\mathbf{a}, \mathbf{b}\big)$.

Corollary 3.2. *Suppose* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ *with* $F_0(\mathbf{a}) = 0$, $F_0(\mathbf{b}-) = 1$ *and* $\mu_{F_0}(\big[\mathbf{a}, \mathbf{b}\big]) = 1$ *satisfy one of the assumptions listed in Proposition* 3.3. *If* $\mu(\{\mathbf{y}\}) > 0$ *for each* $\mathbf{y} \in \big[\mathbf{a}, \mathbf{b}\big] \setminus \mathcal{C}_{F_0}$, *then for each* $\omega \in \Omega_\mu$, $\hat{F}_n(\mathbf{x}; \omega) \to F_0(\mathbf{x})$ *for all* $\mathbf{x} \in \mathbb{R}^2$.

Corollary 3.3. *If every* $\mathbf{y} \in \mathcal{I}_{F_0}$ *is strongly regular with respect to* μ *and* $\mu(\{\mathbf{y}\}) > 0$ *for each* $\mathbf{y} \notin \mathcal{C}_{F_0}$, *then for each* $\omega \in \Omega_\mu$, $\hat{F}_n(\mathbf{x}; \omega) \to F_0(\mathbf{x})$ *for all* $\mathbf{x} \in \mathbb{R}^2$.

We now state propositions on the uniform convergence on the entire $\mathbb{R}^2$ plane and on a closed rectangle.

Proposition 3.5. *Suppose* F_0 *is continuous. If for all* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$, $\mu_{F_0}(\big(\mathbf{a}, \mathbf{b}\big)) > 0$ *implies* $\mu(\big(\mathbf{a}, \mathbf{b}\big)) > 0$, *then the GMLE is uniformly strongly consistent*, i.e.,

$$\sup_{\mathbf{x} \in \mathbb{R}^2} |\hat{F}_n(\mathbf{x}) - F_0(\mathbf{x})| \to 0 \; a.s.$$

Proposition 3.6. *Let* $\mathbf{s}, \mathbf{t} \in \mathbb{R}^2$ *be such that* $\mathbf{s} \leqslant \mathbf{t}$ *and* $\mathbf{s} \neq \mathbf{t}$. *If* $\mathbf{s}$ *and* $\mathbf{t}$ *satisfy the following conditions*:

(a) *either* $\mu(\{\mathbf{s}\}) > 0$ *or* $F_0(\mathbf{s}) = 0$,
(b) *either* $\mu(\{\mathbf{t}\}) > 0$ *or* $F_0(\mathbf{t}-) = 1$,
(c) F_0 *is continuous on* $\big[\mathbf{s}, \mathbf{t}\big]$, *and*
(d) *for all* $\mathbf{a}, \mathbf{b} \in \big[\mathbf{s}, \mathbf{t}\big]$, $\mu_{F_0}(\big(\mathbf{a}, \mathbf{b}\big)) > 0$ *implies* $\mu(\big(\mathbf{a}, \mathbf{b}\big)) > 0$,

then the GMLE is uniformly strongly consistent on $\big[\mathbf{s}, \mathbf{t}\big]$, i.e.,

$$\sup_{\mathbf{x} \in [\mathbf{s}, \mathbf{t}]} |\hat{F}_n(\mathbf{x}) - F_0(\mathbf{x})| \to 0 \; a.s.$$

Uniform consistency results for the GMLE without conditions (a) and (b) have been mistakenly claimed in the literature in the univariate interval censorship models (see [8]). We shall demonstrate clearly in Section 4 that both of these conditions are needed for the proof.

Remark 3.1. For simplicity, the results in Section 3 are stated and proved in the bivariate case. We should point out that the following statements are still valid under the multivariate case, except for Proposition 3.3 and Statements (i) and (ii) of Proposition 3.2. The modification of the latter statements are quite tedious and we shall ignore them. One of the referee of our manuscript points out that the similar versions of Corollary 2.1, Propositions 3.1, 3.4 and 3.5 in this paper were obtained for the bivariate current status data by Song [9].

4. Proofs of propositions

Let $\mathbb{Q}^2$ be the set of all points in $\mathbb{R}^2$ whose coordinates are rational. Then for each $\omega \in \Omega$, there exists a subsequence $\{n'\}$ of $\{n\}$ tending to infinity such that $\hat{F}_{n'}(\mathbf{x};\omega) \to F(\mathbf{x};\omega)$ for all $\mathbf{x} \in \mathbb{Q}^2$, where $F \in \mathcal{F}$. To uniquely determine the F, for each $\mathbf{x} \in \mathbb{R}^2 \setminus \mathbb{Q}^2$, define $F_\omega(\mathbf{x}) = F(\mathbf{x};\omega) = \inf\{F(\mathbf{a};\omega) : \mathbf{a} \in \mathbb{Q}^2 \text{ and } \mathbf{x} \leqslant \mathbf{a}\}$. Since $\hat{F}_n(\cdot;\omega)$ is a distribution function for each n and each ω, F_ω is nondecreasing in each variable and bounded by 0 and 1. For convenience, abbreviate $\hat{F}_n(\cdot;\omega)$ by F_n, and F_ω by F. By Theorem 2.1, $\lim_{n\to\infty} \int |F_n - F_0| d\mu = \int |F - F_0| d\mu = 0$ a.s.. Let $\mathcal{D} = \{\mathbf{x} \in \mathbb{R}^2 : F(\mathbf{x}) \neq F_0(\mathbf{x})\}$. Then, $\mu(\mathcal{D}) = 0$.

Proof of Proposition 3.1. We shall show that if $\mathbf{x}_0 \in \mathcal{D}$ is a continuity point of F_0, then $\mathbf{x}_0$ is not regular. If $\mathcal{C}_{F_0} \cap \mathcal{D} \neq \emptyset$, there exists $\mathbf{x}_0 \in \mathcal{C}_{F_0} \cap \mathcal{D}$ such that $|F(\mathbf{x}_0) - F_0(\mathbf{x}_0)| = d > 0$. Suppose $F(\mathbf{x}_0) > F_0(\mathbf{x}_0)$. Since F_0 is continuous and monotone, there is a $\delta > 0$ such that $|F_0(\mathbf{x}) - F_0(\mathbf{x}_0)| < \frac{d}{2}$ for all $\mathbf{x} \in \mathcal{Q}_\delta[\mathbf{x}_0)$. Furthermore, $|F(\mathbf{x}) - F_0(\mathbf{x})| \geqslant |F(\mathbf{x}_0) - F_0(\mathbf{x})| \geqslant |F(\mathbf{x}_0) - F_0(\mathbf{x}_0)| - |F_0(\mathbf{x}) - F_0(\mathbf{x}_0)| > \frac{d}{2}$, for all $\mathbf{x} \in \mathcal{Q}_\delta[\mathbf{x}_0)$ by monotone property of F. Then $\mathcal{Q}_\delta[\mathbf{x}_0) \subset \mathcal{D}$ with μ-measure 0, i.e., $\mathbf{x}_0$ is not regular. Similarly, if $F(\mathbf{x}_0) < F_0(\mathbf{x}_0)$, then there is a $\delta' > 0$ such that $|F_0(\mathbf{x}_0) - F_0(\mathbf{x})| < \frac{d}{2}$ for all $\mathbf{x} \in \mathcal{Q}_{-\delta'}(\mathbf{x}_0]$. Thus $\mathcal{Q}_{-\delta'}(\mathbf{x}_0]$ is in $\mathcal{D}$ with μ-measure 0, i.e., $\mathbf{x}_0$ is not regular. $\square$

Proof of Proposition 3.2. We shall show that if one of the assumptions is satisfied and $\mathcal{D}$ contains a continuity point of F_0 in $\big(\mathbf{a},\mathbf{b}\big)$, then $\mu(\mathcal{D}) > 0$, which contradicts Theorem 2.1. Let $\mathcal{D}_1 = \mathcal{D} \cap \big(\mathbf{a},\mathbf{b}\big)$.

(i) Assume $a_1 = b_1$ and $\big(\mathbf{a},\mathbf{b}\big) \subset \mathcal{S}2_\mu$. Then $\mathbf{x}_0 \in \mathcal{C}2_{F_0} \cap \mathcal{D}_1$ implies that either $\Big|_{-\delta}(\mathbf{x}_0]$ or $\Big|_{\delta}[\mathbf{x}_0)$ is contained in $\mathcal{D}$ for some positive δ. Since $\mathbf{x}_0 \in \big(\mathbf{a},\mathbf{b}\big) \subset \mathcal{S}2_\mu$, both $\Big|_{\delta}(\mathbf{x}_0)$ and $\Big|_{-\delta}(\mathbf{x}_0)$ have positive μ-measure, which leads to $\mu(\mathcal{D}_1) > 0$.

(ii) Proof is similar to that given above for part (i).

(iii) Assume $\mathbf{a} < \mathbf{b}$ and $\big(\mathbf{a},\mathbf{b}\big) \subset \mathcal{S}_\mu$. Let $\mathbf{x}_0 \in \mathcal{C}_{F_0} \cap \mathcal{D}_1$, say $|F(\mathbf{x}_0) - F_0(\mathbf{x}_0)| > 0$. Since F and F_0 are both monotone and $\big(\mathbf{a},\mathbf{b}\big) \subset \mathcal{S}_\mu$. $\mathbf{x}_0$ is a continuity point of F_0, there is a $\delta > 0$ such that either $\mathcal{Q}_{-\delta}(\mathbf{x}_0]$ or $\mathcal{Q}_\delta[\mathbf{x}_0)$ is contained in $\mathcal{D}$. Since $\mathbf{x}_0$ is an interior point of $\mathcal{S}_\mu$, both $\mathcal{Q}_\delta(\mathbf{x}_0)$ and $\mathcal{Q}_{-\delta}(\mathbf{x}_0)$ have positive μ-measure. This implies $\mu(\mathcal{D}_1) > 0$. $\square$

Proof of Proposition 3.3. Suppose $F_0(\mathbf{a}) = 0$, $F_0(\mathbf{b}-) = 1$ and $\mu_{F_0}(\big[\mathbf{a}, \mathbf{b}\big]) = 1$, for some $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ such that $\mathbf{a} \leqslant \mathbf{b}$. Let $\mathcal{D}_1 = \big[\mathbf{a}, \mathbf{b}\big] \cap \mathcal{D}$.

(i) Let $a_1 = b_1$ and $\big(\mathbf{a}, \mathbf{b}\big) \subset \mathcal{S}2_\mu$. Note that $\mathbf{a}, \mathbf{b} \in \mathcal{C}2_{F_0}$. Then $\mathbf{a} \notin \mathcal{D}_1$, otherwise there is a $\delta > 0$ such that $\big|_{\delta}(\mathbf{a}) \subset \mathcal{D}_1$, and thus $\mathcal{D}_1$ has positive μ-measure, a contradiction. Also, $\mathbf{b} \notin \mathcal{D}_1$, otherwise there is a $\big|_{-\delta}(\mathbf{b}) \subset \mathcal{D}_1$, and thus leads to the contradiction $\mu(\mathcal{D}_1) > 0$. In view of Proposition 3.2, $F(\mathbf{x}) = F_0(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{C}2_{F_0} \cap \big(\mathbf{a}, \mathbf{b}\big)$. Since $\mu_F([a_1, a_1] \times [a_2, b_2]) = 1$ (as $a_1 = b_1$) and $\mu_{F_0}([a_1, a_1] \times [a_2, b_2]) = 1$, $\mu_F((a_1, a_1 + \delta] \times [a_2, b_2]) = 0$ and $\mu_{F_0}((a_1, a_1 + \delta] \times [a_2, b_2]) = 0$. This implies that for each $\mathbf{x} \in [a_1, a_1] \times [a_2, b_2]$, $F(\mathbf{y}) = F(\mathbf{x})$ and $F_0(\mathbf{y}) = F_0(\mathbf{x})$, where $\mathbf{y} \in (x_1, x_1 + \delta) \times [x_2, x_2]$ $(\delta > 0)$. Hence, if $\mathbf{x}$ is a vertical continuity point of F, then it implies that so is $\mathbf{y}$ for all $\mathbf{y} \in (x_1, x_1 + \delta) \times [x_2, x_2]$ $(\delta > 0)$. Verify that $F(\mathbf{x}) = F_0(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{R}^2 \setminus \big[\mathbf{a}, \infty\mathbf{1}\big)$ and $F(\mathbf{x}) = F_0(\mathbf{x}) = 1$ for all $\mathbf{x} \in \big[\mathbf{b}, \infty\mathbf{1}\big)$. Therefore, $F(\mathbf{x}) = F_0(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{C}2_{F_0}$.

(ii) Proof is similar to that given above for part (i).

(iii) $\mathbf{a} < \mathbf{b}$ and $\big(\mathbf{a}, \mathbf{b}\big) \subset \mathcal{S}_\mu$. Note that $\mathbf{a}, \mathbf{b} \in \mathcal{C}_{F_0}$. Thus $\mathbf{a} \notin \mathcal{D}_1$, otherwise there is $\mathcal{Q}_\delta(\mathbf{a}) \subset \mathcal{D}_1$, and thus $\mu(\mathcal{D}_1) > 0$, a contradiction. Similarly, $\mathbf{b} \notin \mathcal{D}_1$. Notice that $\mu_{F_0}(\big[\mathbf{a}, \mathbf{b}\big]) = F_0(\mathbf{b}) + F_0(\mathbf{a}-) - F_0(a_1-, b_2) - F_0(b_1, a_2-)$. For each $\mathbf{x} \in ((-\infty, a_1) \times [a_2, b_2]) \cup ([a_1, b_1] \times (-\infty, a_2))$, $F_0(\mathbf{x}) = 0$. Consequently, $\hat{F}_n(\mathbf{x}) = 0$ and thus $F(\mathbf{x}) = F_0(\mathbf{x}) = 0$. Moreover, similar to part (i), we establish

1. $F(\mathbf{x}) = F_0(\mathbf{x})$ for all $\mathbf{x} \in C1_{F_0} \cap (\big[\binom{a_1}{b_2}, \binom{b_1}{\infty}\big] \cup \partial_b\big[\mathbf{a}, \mathbf{b}\big])$ and for all $\mathbf{x} \in C2_{F_0} \cap (\big[\binom{b_1}{a_2}, \binom{\infty}{b_2}\big] \cup \partial_l\big[\mathbf{a}, \mathbf{b}\big])$;
2. $F(\mathbf{x}) = F_0(\mathbf{x})$ for all $\mathbf{x} \in \big(\mathbf{a}, \mathbf{b}\big) \cap \mathcal{C}_{F_0}$ (by Proposition 3.2);
3. $F(\mathbf{x}) = F_0(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{R}^2 \setminus \big[\mathbf{a}, \infty\mathbf{1}\big)$;
4. $F(\mathbf{x}) = F_0(\mathbf{x}) = 1$ for all $\mathbf{x} \in \big[\mathbf{b}, \infty\mathbf{1}\big)$.

Thus, $F(\mathbf{x}) = F_0(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{C}1_{F_0} \cup \mathcal{C}2_{F_0}$. $\square$

Proof of Proposition 3.4. Let $\mathbf{x}_0 \in \mathcal{C}_{F_0}$. If $\mathbf{x}_0 \in \mathcal{I}_{F_0}$, then $\mathbf{x}_0$ is strongly regular by assumption in the proposition, and hence not in $\mathcal{D}$ by Proposition 3.1. Now, suppose $\mathbf{x}_0 \notin \mathcal{I}_{F_0}$. We shall show that $\mathbf{x}_0 \notin \mathcal{D}$. Otherwise, $|F(\mathbf{x}_0) - F_0(\mathbf{x}_0)| > 0$. If $F(\mathbf{x}_0) > F_0(\mathbf{x}_0)$, let $\mathbf{x} = \sup\{\mathbf{x}_0 + \delta\mathbf{1} : F_0(\mathbf{x}_0 + \delta\mathbf{1}) = F_0(\mathbf{x}_0),\ \ \delta > 0\}$. Then $\mathbf{x} \in \mathcal{I}_{F_0}$ and $\mathbf{x} = \mathbf{x}_0 + \delta_0\mathbf{1}$ for some $\delta_0 > 0$. Thus $\mu(\mathcal{Q}_{-\delta_0}(\mathbf{x})) > 0$ by the assumption that $\mathbf{x}$ is strongly regular as $\mathbf{x} \in \mathcal{I}_{F_0}$. Since $F(\mathbf{x}-) \geqslant F(\mathbf{x}_0) > F_0(\mathbf{x}_0) = F_0(\mathbf{x}-)$, $\mathcal{Q}_{-\delta_0}(\mathbf{x}) \subset \mathcal{D}$, which implies that $\mu(\mathcal{D}) > 0$, a contradiction. On the other hand, if $F(\mathbf{x}_0) < F_0(\mathbf{x}_0)$, let $\mathbf{x} = \inf\{\mathbf{x}_0 - \delta\mathbf{1} : F_0(\mathbf{x}_0 - \delta\mathbf{1}) = F_0(\mathbf{x}_0),\ \ \delta > 0\}$. Then $\mathbf{x} \in \mathcal{I}_{F_0}$, $F(\mathbf{x}) \leqslant F(\mathbf{x}_0) < F_0(\mathbf{x}_0) = F_0(\mathbf{x})$, $\mathcal{Q}_{\delta_0}[\mathbf{x}) \subset \mathcal{D}$ for some $\delta_0 > 0$. Consequently, $\mu(\mathcal{D}) > 0$, a contradiction again. $\square$

Proof of Proposition 3.5. We shall show $\mathcal{D} = \emptyset$. Otherwise, let $\mathbf{x}_0 \in \mathcal{D}$. If $F(\mathbf{x}_0) - F_0(\mathbf{x}_0) = d > 0$, let $\mathbf{x} = \sup\{\mathbf{x}_0 + \delta\mathbf{1} : F_0\big(\mathbf{x}_0 + \delta\mathbf{1}\big) = F_0(\mathbf{x}_0),\ \ \delta > 0\}$. Then $\mathbf{x} \in \mathcal{I}_{F_0}$. Since F_0 is continuous, there is a positive δ_0 such that $F_0(\mathbf{x} + \delta_0\mathbf{1}) - F_0(\mathbf{x}) < \frac{d}{2}$. Then $\mu_{F_0}(\big(\mathbf{x}, \mathbf{x} + \delta_0\mathbf{1}\big)) > 0$ and $\big(\mathbf{x}, \mathbf{x} + \delta_0\mathbf{1}\big) \subset \mathcal{D}$, which imply that $\mu(\mathcal{D}) > 0$, a contradiction to Theorem 2.1. The same contradiction can be reached similarly for the case $F(\mathbf{x}_0) - F_0(\mathbf{x}_0) < 0$. Thus $\mathcal{D} = \emptyset$ and F_n converges to F_0 pointwise.

Let $\varepsilon > 0$ and $\mathbf{x}_0 \in \mathbb{R}^2$. By continuity and monotonicity of F_0, we can choose finitely many quantiles $\{a_0, a_1, \ldots, a_\alpha\}$ and $\{b_0, b_1, \ldots, b_\beta\}$ such that $a_0 = b_0 = -\infty$, $F_0(a_i, \infty) - F_0(a_{i-1}, \infty) < \varepsilon/2$ for each $i = 1, \ldots, \alpha$, and $F_0(\infty, b_j) - F_0(\infty, b_{j-1}) < \varepsilon/2$ for each $j = 1, \ldots, \beta$. Then there exists an N such that $|F_n(a_i, b_j) - F_0(a_i, b_j)| < \varepsilon$ for all $i = 0, \ldots, \alpha$, $j = 0, \ldots, \beta$, and all $n > N$. Recall that we denote the product set $(a_i, a_{i+1}] \times (b_j, b_{j+1}]$ by $\left(\binom{a_i}{b_j}, \binom{a_{i+1}}{b_{j+1}}\right]$, and denote $\left(\binom{a_i}{b_j}, \binom{a_{i+1}}{b_{j+1}}\right)$ in a similar manner. Since $\left(\binom{a_i}{b_j}, \binom{a_{i+1}}{b_{j+1}}\right]$'s are a partition of $\mathbb{R}^2$, $\mathbf{x}_0 \in \left(\binom{a_i}{b_j}, \binom{a_{i+1}}{b_{j+1}}\right]$ for some i, j. Then $|F_0(\mathbf{y}) - F_0(\mathbf{x}_0)| < \varepsilon$ for all $\mathbf{y} \in \left(\binom{a_i}{b_j}, \binom{a_{i+1}}{b_{j+1}}\right]$. Moreover, by the pointwise convergence of F_n to F_0, there exists an N such that for all $n > N$, we have

$$\sup_{\mathbf{x}\in\mathbb{R}^2} |F_n(\mathbf{x}) - F_0(\mathbf{x})| \leqslant \sup_{i,j} \sup_{\mathbf{x}\in(a_i,a_{i+1}]\times(b_j,b_{j+1}]} |F_n(\mathbf{x}) - F_0(\mathbf{x})|.$$

$$\begin{aligned}
&\sup_{\mathbf{x}\in(a_i,a_{i+1}]\times(b_j,b_{j+1}]} F_n(\mathbf{x}) - F_0(\mathbf{x}) \\
&\leqslant F_n(a_{i+1}, b_{j+1}) - F_0(a_i, b_j) \\
&= F_n(a_{i+1}, b_{j+1}) - F_0(a_{i+1}, b_{j+1}) + F_0(a_{i+1}, b_{j+1}) - F_0(a_i, b_j) \\
&\leqslant \varepsilon + 2\varepsilon = 3\varepsilon.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\sup_{\mathbf{x}\in(a_i,a_{i+1}]\times(b_j,b_{j+1}]} F_n(\mathbf{x}) - F_0(\mathbf{x}) &\geqslant F_n(a_i, b_j) - F_0(a_{i+1}, b_{j+1}) \\
&= F_n(a_i, b_j) - F_0(a_i, b_j) + F_0(a_i, b_j) \\
&\quad - F_0(a_{i+1}, b_{j+1}) \\
&\geqslant -\varepsilon - 2\varepsilon = -3\varepsilon.
\end{aligned}$$

Thus

$$\sup_{\mathbf{x}\in\mathbb{R}^2} |F_n(\mathbf{x}) - F_0(\mathbf{x})| \leqslant 3\varepsilon. \qquad \square$$

Proof of Proposition 3.6. *WLOG*, assume $\mathbf{s} < \mathbf{t}$. First consider the case $\mu(\{\mathbf{s}\}) > 0$ and $F_0(\mathbf{t}-) = 1$. By Corollary 2.1, $F(\mathbf{s}) = F_0(\mathbf{s})$. If $F_0(\mathbf{s}) = 1$, the proof is completed. $\square$

Consider $F_0(\mathbf{s}) < 1$. We shall show $\mathcal{D}_1 = \mathcal{D} \cap [\mathbf{s}, \mathbf{t}] = \emptyset$ in three steps.

(1) $\mathbf{t} \notin \mathcal{D}_1$. Otherwise, $F_0(\mathbf{t}) - F(\mathbf{t}) = d > 0$ as $F_0(\mathbf{t}-) = 1$. Since $F_0(\mathbf{s}) < 1$, if we let $\mathbf{x} = \inf\{\delta\mathbf{t} + (1-\delta)\mathbf{s} : F_0(\mathbf{t}) = F_0(\delta\mathbf{t} + (1-\delta)\mathbf{s}), \delta > 0\}$, then $\mathbf{x}$ is either $\mathbf{t}$ or a member of $(\mathbf{s}, \mathbf{t})$. Also, $\mathbf{x} \in \mathcal{I}_{F_0}$. By continuity of F_0, if $\mathbf{x} = \mathbf{t}$, then for some $\delta > 0$, $\delta\mathbf{t} + (1-\delta)\mathbf{s} \in (\mathbf{s}, \mathbf{t})$ and $0 < F_0(\mathbf{t}) - F_0(\delta\mathbf{t} + (1-\delta)\mathbf{s}) < \frac{d}{2}$, which implies that $\mathcal{D}_1$ contains $\mathcal{Q}_{-\delta}(\mathbf{t})$ with positive μ-measure, a contradiction. If $\mathbf{x} \in (\mathbf{s}, \mathbf{t})$, there also exits a $\delta > 0$ such that $|F_0(\mathbf{y}) - F_0(\mathbf{t})| < \frac{d}{2}$ for all $\mathbf{y} \in (\mathbf{x} - \delta\mathbf{1}, \mathbf{x}) \subset (\mathbf{s}, \mathbf{t})$, and thus $\mathcal{G}_\delta(\mathbf{x})$ with positive μ-measure is in $\mathcal{D}_1$, also a contradiction.

(2) $\partial[\mathbf{s}, \mathbf{t}] \cap \mathcal{D}_1 = \emptyset$. Otherwise, let $\mathbf{x}_0 \in \mathcal{D}_1 \cap \partial[\mathbf{s}, \mathbf{t}]$.

Suppose $\mathbf{x}_0 \in \mathcal{D}_1 \cap \partial_u[\mathbf{s}, \mathbf{t}]$. If $F(\mathbf{x}_0) > F_0(\mathbf{x}_0)$, let $\mathbf{x} = \sup\{\mathbf{x}_0 + \binom{\delta}{0} : F_0(\mathbf{x}_0 + \binom{\delta}{0}) = F_0(\mathbf{x}_0), \delta > 0\}$. The continuity of F_0 implies that $\mathbf{x} \in \partial_u[\mathbf{s}, \mathbf{t}] \setminus \{\mathbf{t}\}$ and $F_0(\mathbf{x}) < 1$. This fact together with condition (d) imply that there exists a $\delta > 0$ such that $\mathrel{-\!\!\cdot\!\!-}_{\delta}(\mathbf{x})$ has positive μ-measure and is a subset of $\mathcal{D}_1$, a contradiction.

Now assume $F_0(\mathbf{x}_0) > F(\mathbf{x}_0)$. Let $\mathbf{x} = \inf\{\mathbf{x}_0 - \binom{\delta}{0} : F_0(\mathbf{x}_0 - \binom{\delta}{0}) = F_0(\mathbf{x}_0), \delta > 0\}$. Then either $\mathbf{x} = \binom{s_1}{t_2}$ or $\mathbf{x} \in \partial_u\big[\mathbf{s}, \mathbf{t}\big] \setminus \{\binom{s_1}{t_2}\}$. In the first case, there exists a $\delta > 0$ such that $0 < F_0(\mathbf{x}) - F_0(\mathbf{x} - \binom{0}{\delta}) < \frac{d}{2}$ and $\dot{|}_{-\delta}(\mathbf{x})$ is contained in $\partial_l\big[\mathbf{s}, \mathbf{t}\big] \setminus \{\mathbf{s}\}$ since $F_0(\mathbf{x}_0) > F(\mathbf{x}_0) \geqslant F(\mathbf{s}) = F_0(\mathbf{s})$. Then $\dot{|}_{-\delta}(\mathbf{x})$ has positive μ-measure and is contained in $\mathcal{D}_1$. In the second case, there exists a subset of $\mathcal{D}_1$ with positive μ-measure, namely, $\mathrel{-\!\!\cdot\!\!-}_{-\delta}(\mathbf{x})$ for some $\delta > 0$, a contradiction.

Similarly, if $\mathbf{x}_0 \in \mathcal{D}_1$ is contained in the boundary of $\big[\mathbf{s}, \mathbf{t}\big]$, namely, $\partial_l\big[\mathbf{s}, \mathbf{t}\big]$, $\partial_r\big[\mathbf{s}, \mathbf{t}\big]$ and $\partial_b\big[\mathbf{s}, \mathbf{t}\big]$, the same contradiction is reached. Hence $\partial\big[\mathbf{s}, \mathbf{t}\big] \setminus \{\mathbf{t}\}$ is not in $\mathcal{D}_1$.

(3) In view of the first part in the proof of Proposition 3.5, $\mathcal{D}_1 \cap \big(\mathbf{s}, \mathbf{t}\big) = \emptyset$, otherwise, we can find an $\mathbf{x} \in \mathcal{I}_{F_0}$ such that $\mathbf{x} \in \big(\mathbf{s}, \mathbf{t}\big)$ and construct an open square around it with positive μ-measure that is also contained in $\mathcal{D}_1$.

Note that conditions (a) and (b) in Proposition 3.6 imply the following four combinations: (I) $\mu(\{\mathbf{s}\}) > 0$ and $F_0(\mathbf{t}-) = 1$, (II) $F_0(\mathbf{s}) = 0$ and $\mu(\{\mathbf{t}\}) > 0$, (III) $\mu(\{\mathbf{s}\}) > 0$ and $\mu(\{\mathbf{t}\}) > 0$, (IV) $F_0(\mathbf{s}) = 0$ and $F_0(\mathbf{t}-) = 1$. We have just proved the proposition in case (I). Similar proof will go through for the other three cases.

Now, we have shown that F_n converges pointwise to F_0 in $\big[\mathbf{s}, \mathbf{t}\big]$. By assumption, F_0 is continuous on the bounded closed set $\big[\mathbf{s}, \mathbf{t}\big]$. Let $\varepsilon > 0$. Similar to the second part in the proof of Proposition 3.5, we can select finitely many quantiles $\{a_0, a_1, \ldots, a_\alpha\}$ such that $a_0 = s_1$, $a_\alpha = t_1$ and $F_0(a_i, t_2) - F_0(a_{i-1}, t_2) < \varepsilon$ for all $i = 1, \ldots, \alpha$, and quantiles $\{b_0, b_1, \ldots, b_\beta\}$ such that $b_0 = s_2$, $b_\beta = t_2$ and $F_0(t_1, b_j) - F_0(t_1, b_{j-1}) < \varepsilon$ for all $j = 1, \ldots, \beta$. Then there exist an $N > 0$ satisfying $|F_n(a_i, b_j) - F_0(a_i, b_j)| < \varepsilon$ for all $n > N$, and for all $i = 0, \ldots, \alpha$ and $j = 0, \ldots, \beta$. For each $\mathbf{x}_0 \in \big[\mathbf{s}, \mathbf{t}\big]$, $\mathbf{x}_0 \in \big[\binom{a_i}{b_j}, \binom{a_{i+1}}{b_{j+1}}\big]$ for some i, j. Thus, $|F_n(\mathbf{x}_0) - F_m(\mathbf{x}_0)| \leqslant 12\varepsilon$ for all $n, m > N$ by an argument similar to that used in the inequality (4.1). Since ε is arbitrary, we obtain the uniform convergence of the GMLE in the closed rectangle $\big[\mathbf{s}, \mathbf{t}\big]$. $\square$

Acknowledgements

Thanks go to the editor and two referees for valuable suggestions. Q.Q. Yu and G.Y.C Wong were partially supported by DOD Grant DAMD17-99-1-9390.

References

[1] M. Ayer, H.D. Brunk, G.M. Ewing, W.T. Reid, E. Silverman, An empirical distribution function for sampling incomplete information, Ann. Math. Statist. 26 (1955) 641–647.

[2] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.

[3] R. Gentleman, C.J. Geyer, Maximum likelihood for interval censored data: consistency and computation, Biometrika 81 (1994) 618–623.

[4] P. Groeneboom, J.A. Wellner, Information Bounds and Nonparametric Maximum Likelihood Estimation, Birkhäuser Verlag, Basel, 1992.

[5] J. Huang, Efficient estimation for proportional hazards models with interval censoring, Ann. Statist. 24 (1996) 540–568.

[6] J.J. Ren, Goodness of fit tests with interval censored data, Scanad. J. Statist. 30 (2003) 211–226.
[7] W. Rudin, Principles of Mathematical Analysis, McGraw-Hill, New York, 1976.
[8] A. Schick, Q.Q. Yu, Consistency of the GMLE with mixed case interval-censored data, Scanad. J. Statist. 27 (2000) 45–55.
[9] S. Song, Estimation with bivariate interval censored data, Unpublished Ph.D Dissertation, University of Washington, 2001.
[10] S. Song, Estimation with univariate "mix case" interval censored data, Statistica Sinica 14 (2004) 269–282.
[11] J.G. Sun, H.B. Fang, A nonparametric test for panel count data, Biometrika 90 (2003) 199–208.
[12] The Italian–American Cataract Study Group, Incidence and progression of cortical, nuclear, and posterior subcapsular cataracts, Amer. J. Ophthal. 118 (1994) 623–631.
[13] A. van der Vaart, J.A. Wellner, Preservation theorems for Glivenko–Cantelli and uniform Glivenko–Cantelli class, in: E. Gine, D.M. Mason, J.A. Wellner (Eds.), High Dimensional Probability, vol. II, Birkhäuser, Boston, 2000, pp. 115–133.
[14] J.A. Wellner, Interval censoring case 2: alternative hypotheses. in: H.L. Koul, J.V. Deshpande (Eds.), Analysis of Censored Data, Proceedings of the Workshop on Analysis of Censored Data, University of Pune, Pune, India, IMS Lecture Notes, Monograph Series, vol. 27, 1995, pp. 271–291, December 28, 1994–January 1, 1995.
[15] J.A. Wellner, Y. Zhang, Two estimators of the mean of a counting process with panel count data, Ann Statist. 28 (2000) 779–814.
[16] G.Y.C. Wong, Q.Q. Yu, Generalized MLE of a joint distribution function with multivariate interval-censored data, J. Multivariate Anal. 69 (1999) 155–166.
[17] S.H. Yu, Consistency of the generalized MLE with multivariate mixed case interval-censored data, Ph.D Dissertation, Binghamton University, 2000.
[18] Q.Q. Yu, L.X. Li, On the strong consistency of the product limit estimator, Sankhya A 56 (1994) 416–430.
[19] Q.Q. Yu, A. Schick, L.X. Li, G.Y.C. Wong, Asymptotic properties of the GMLE with case 2 interval-censored data, Statist. Probab. Lett. 37 (1998) 223–228.
[20] Q.Q. Yu, G.Y.C. Wong, Q.M. He, Estimation of a joint distribution function with multivariate interval-censored data when the nonparametric MLE is not unique, Biometrical J. 42 (2000) 747–763.
[21] Y. Zhang, W. Liu, Y.H. Zhan, A nonparametric two-sample test of the failure function with interval censoring case 2, Biometrika 88 (2001) 677–686.